

Saccadic Suppression During Reading Activity:  
Is the Spill-Over Effect Weaker After a Longer Saccade?

by  
Kiyomi Yatabe

A thesis submitted in partial fulfillment  
of the requirements for the degree  
of  
Master of Science  
in  
Psycholinguistics

UNIVERSITY OF EDINBURGH  
Edinburgh, UK  
August, 2006

**Abstract**

Although it has generally been thought that the duration of saccades should be subtracted from the reading time in eye movement research, Irwin (1998) has demonstrated that lexical processing such as word recognition is not suppressed during saccades, and has thus called that conventional wisdom into question (Murray, 2000; Radach & Kennedy, 2004; Rayner, 1998; Vonk & Cozijn, 2003). In Irwin's experiments, however, the subjects were merely reading isolated words, not sentences. It is therefore not clear what we can legitimately infer from the results of these experiments about what happens during saccades when people are reading sentences. In this thesis, I report the result of an eye-tracking experiment which tests whether cognitive processing for reading continues during saccades when people are reading sentences. The subjects in the experiment read the first half of a sentence, and then made either a short or a long saccade, before going on to read the rest of the sentence; in half of the experimental sentences, a low-frequency word was used at the location immediately preceding the saccade, so that some cognitive processing regarding that word would "spill over," that is, continue even after the eyes have left the word. The result of the experiment showed that the spill-over effect, that is, the difference between the effect that a high-frequency word has on the processing of the ensuing material and the effect that low-frequency word has on the processing of the ensuing material, was smaller after a longer saccade. This result indicates that linguistic processing is not completely suppressed during saccades when people are reading sentences.

## Contents

1	Introduction .....	4
1.1.	Cognitive suppression during saccades and eye movement research .....	4
1.2.	Lexical Processing Experiments by Irwin (1998) .....	5
1.3.	Limitations of Irwin's experiments .....	6
1.3.1.	Materials .....	6
1.3.2.	Independent variables .....	8
1.3.3.	Dependent variables .....	9
2	A sketch of our experiment .....	10
2.1.	Materials .....	10
2.2.	Independent variables .....	10
2.3.	Dependent variables .....	11
3	Method .....	12
3.1.	Subjects .....	12
3.2.	Materials .....	12
3.3.	Assignment of the materials .....	14
3.4.	Apparatus .....	15
3.5.	Procedure .....	16
3.6.	Data Analyses .....	18
4	Results .....	21
4.1.	Ancillary results: how the subjects looked at the target words themselves .....	21
4.2.	Primary results: Spill-over effects after the saccade.....	23
5	Discussion .....	28
5.1.	Interpretations of results .....	28
5.2.	Future directions .....	33
5.2.1.	Leftward saccades .....	33
5.2.2.	Other types of linguistic material .....	34
6	Summary .....	34
7	Acknowledgments .....	35
8	References .....	35
	Appendix .....	41

## 1. Introduction

The aim of this thesis is to study cognitive processing that takes place during saccades when people are reading a sentence or a sequence of sentences, as opposed to when people are reading isolated words. Specifically, the present study is concerned with the way people move their eyes while they read sentences, and what it reveals about the underlying cognitive events.

I will first describe the background of this study in this section and sketch the design of the experiment conducted in the present study in section 2. Then I will present the details of the design of the experiment in section 3, report the results of the experiment in section 4, and examine what they mean in section 5.

In this introductory section, I will survey the previous literature that forms the background of the present study. After presenting a broad background in subsection 1.1, I will provide, in subsection 1.2, a brief exposition of Irwin's (1998) experiments, which constitute the point of departure for the present study. Then in Section 1.3, I will discuss some considerations that suggest that the results of Irwin's experiments might not necessarily give us a realistic picture of what happens when people read sentences.

### 1.1. Cognitive suppression during saccades and eye movement research.

The foveal high acuity vision corresponding to the center of our gaze is confined to a relatively small retinal area encompassing approximately 3 degrees of central vision. This anatomical limitation of our visual perception requires *saccades*, ballistic jerky eye movements from point to point in space for fine details to be resolved. Saccades occur several times a second (every 250-300 ms) and generally have duration of between 30 and 100 ms during reading, depending on the distance the movement covers (see e.g. Rayner, 1998).

During a saccadic eye movement, visual sensitivity is greatly reduced, thereby confining the vast majority of visual perception to the fixations between saccades. Because of this phenomenon called saccadic (visual) suppression, we are unaware of blurred motion images produced by saccades and are not disturbed by them. Researchers have been interested in whether higher-level cognitive processing might also be suppressed during saccades.

Potter (1984) reported that people can read and comprehend a sentence faster when saccades are eliminated by presenting text one word at a time at the same location in the visual field in rapid serial visual presentation (RSVP) mode. At the reading rate of 720 words per minute, the performance rate for recall and plausibility judgment in RSVP mode,

in which the subjects were not able to see more than one word at the same time or go back to words that they had seen previously, was better than that in normal text mode in which the same sentence was displayed conventionally for the same total time. This result suggests that higher-level cognitive processing such as those involved in reading activity are suppressed during saccades. In accordance with Potter's finding, the duration of saccades is conventionally subtracted from the reading time in most eye movement experiments; his work had an effect of making this practice, which was already widely followed in the 70s, more entrenched.

On the other hand, Irwin (1998) demonstrated that lexical processing is not suppressed during saccades and concluded that difference in reading speed between RSVP text and normal text is not due to suppression of cognitive processing during saccades but due to other consequences of eye movements, such as motor planning and visual saccadic suppression (see also, Irwin, 2003). Furthermore, based on the result of the experiment on lexical processing during saccade, Irwin (1998, 2004) proposed that the duration of saccades in a region should not be subtracted from the reading comprehension time associated with that region (see also, Inhoff & Radach, 1998).

This issue, whether duration of saccades in a region should be included to the reading comprehension time associated with that region, remains unresolved among eye movement researchers at the moment (Murray, 2000; Radach & Kennedy, 2004; Rayner, 1998; Vonk & Cozijn, 2003). The aim of our research is to contribute to the resolution of this issue.

## 1.2. *Lexical Processing Experiments by Irwin (1998).*

Irwin's (1998) experiments constitute the point of the departure for the present study.

Irwin (1998) attempted to see if linguistic processing was suppressed during saccades by examining whether people can perform lexical-decision tasks as fast while making a saccade as they can when they are not making a saccade. If linguistic processing is not suppressed at all during saccades, people will be able to perform the tasks just as fast while making a saccade as when they are not making a saccade. This was the central prediction that was tested in this study, although, for reasons that will be noted below, the comparison was in fact made between a condition involving a long saccade and a condition involving a short saccade, rather than between one involving a saccade and one involving no saccade.

The stimuli used in the experiment were words or pronounceable non-words consisting of four letters. Six lists of stimuli were constructed; each list contained 30 four-letter words and 30 four-letter non-words that were constructed by changing one of the letters in a word. The mean written frequency of the words was roughly the same in all the lists.

In an experimental session, first a fixation box appeared on the left side of the display. The subject fixated the center of this box for 1500 ms, and then a letter string was presented inside it. A saccade target box was presented on the right side of the display either 7.5 degrees or 40 degrees away from the letter string, at the same time as the letter string was presented on the left side of the display. The subjects were instructed to make a saccade to the target box, to decide whether the letter string was a word or non-word while making a saccade to the target box, and to respond as quickly but as accurately as possible. The letter string was erased when the eye-tracker detected saccade onset. Subjects pressed a button held in their dominant hand to indicate that the letter string was a word and a button held in their non-dominant hand to indicate that it was a non-word.

If the process of word recognition is suppressed during saccades, the response time, the time that elapsed between the onset of letters and the subject's response, should be longer when subjects have to execute a longer saccade. Since the saccade duration for 40 degrees saccades was about 94 ms and that for 7.5 degrees saccades was about 28 ms, the response time should be about 66 ms longer in the 40 degrees saccade condition than in the 7.5 degrees saccade condition, if the suppression was complete. However, the response time was almost identical regardless of whether the saccade was long or short. This means that the process for word recognition was not suppressed during saccades.

### 1.3. *Limitations of Irwin's experiments.*

In this section, I will discuss some considerations that suggest that the results of Irwin's experiments might not necessarily give us a realistic picture of what happens when people read sentences.

#### 1.3.1. *Materials.*

In the experiments reported in Irwin (1998), the subjects were merely reading isolated words or non-words, not sentences. This is a significant limitation of the experimental design in question, at least if our goal is to obtain information as to what happens when people read normal text, which consists of sentences; it is not clear what we can legitimately infer from the results of these experiments about what happens during saccades when people are reading sentences.

There are several strands of research that shows that what happens when people read sentences is different from what happens when people read isolated words.

Balota, Paul and Spieler (1999) address a question of whether results obtained from the experimental paradigms used in word recognition research may be generalized to normal

reading of sentences. They argue that different pathways are engaged depending on whether the reader is paying attention to isolated words or reading sentences. When attention is directed toward the message conveyed, it appears that readers do not merely rely on locally driven constraints but rather capitalize also on higher level sources such as text integration processes that increase the efficiency of extracting message-level information. To be more specific, predictions that are generated by the preceding context can influence the way the current word is processed. This effect has been well documented (Balota, Pollatsek, & Rayner, 1985; Ehrlich & Rayner, 1981; Kliegl, Nuthmann, & Engbert, 2006; Rayner & Well, 1996).

Furthermore, many studies have shown that perceptual span during reading includes not just the fixated word but the next one or two words as well and that this parafoveal preview has an effect on the way people read (e.g. Murray, 2000; Rayner, 1998; Rayner & Pollatsek, 1989). If the reader fixates the parafoveal word subsequently, he or she uses information obtained prior to the fixation to facilitate processing. This effect is known as parafoveal preview benefit. Likewise, it has been found that the duration of time that a person fixates a word can be affected by the next word, suggesting that some processing of that next word can begin while the eye is fixating the previous word (e.g. Kennedy & Pynte, 2005). Experiments with isolated words are unable to incorporate these effects that parafoveal preview has on the process of reading.

Thus, in order to study what happens when people read sentences, it is necessary to use sentences as experimental materials.

Vonk and Cozijn (2003) recognized this limitation of Irwin's experiments, and did an experiment using sentences as the experimental materials. They measured the length of time it took the subjects to read each region of experimental sentences, and analyzed the results of the experiment in more than one way, including or excluding saccade durations from the reading time. The difference in reading time between hard-to-process regions and easy-to-process regions was greater when saccade durations were included in reading time. On the basis of this result, Vonk and Cozijn suggest that saccade durations should always be included in the measures of reading time. This is a reasonable conclusion, if our goal is to come up with an efficient, reliable way to determine which portions of a given sentence are hard to process and which portions are easy to process. However, their result does not necessarily indicate that linguistic processing continues uninterrupted during saccades and that a measure of reading time that does not exclude the duration of saccades reflects the time needed for processing of a given region more accurately than a measure of reading time that does exclude the duration of saccades. For all we know, the former measure might be artificially inflating the difference in processing difficulty between hard-to-process regions and easy-to-process

regions. We need a different experimental design, if our goal is to see whether linguistic processing continues during saccades or not.

### 1.3.2. *Independent variables.*

The independent variables in Irwin's experiments were saccade length and type of the stimulus (such as a word versus a non-word).

The first variable has the following potential problem. Building on the study reported in van Duren and Sanders (1995), Irwin made comparisons not between saccade and no-saccade conditions, but across saccades of different lengths, in order to assess whether the cognitive suppression occurred during saccades. This method has an advantage of controlling for factors such as motor planning and for adverse perceptual consequences that might accompany saccades. Therefore Irwin's independent variable was saccade length, which was either 7.5 degrees or 40 degrees. One question that arises here concerns the adequacy of the saccade lengths that were used in this experiment. Some researchers have suggested that different degrees of saccades might be subserved by functionally distinct pathways such as dorsal and ventral pathways (Schall, Hanes, Thomson, & King, 1995; Schall, Morel, King, & Bullier, 1995; Stanton, Bruce, & Goldberg, 1993). Saccades of less than 10 degrees direct gaze to salient features in the scene, and are governed by the frontal eye field (FEF), which receives information from the foveal representation in areas such as MT, V4, TEO, TE, and LIP in the parietal cortex area, thus from both the dorsal and ventral pathways. On the other hand, larger amplitude saccades are not primarily guided by object recognition processes, because of the impoverished acuity of peripheral vision; they are instead guided by the peripheral or multimodal visual field representation, thus by the dorsal pathway. Since those two distinct visual pathways are thought to have their own specialized functions (Haxby et al., 1994; Mellet et al., 1996; Mishkin, Ungerleider, & Macko, 1983; Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982), we need to be careful in choosing the degrees of saccades that are used in our experiments. It seems unclear whether saccades of 7.5 degrees can be safely classified as "large" saccades which are executed through the dorsal pathway; there is a possibility that saccades of 7.5 degrees relies on the ventral pathway as well as the dorsal pathway. If such saccades do in fact rely on both the ventral and the dorsal pathway, then the difference between saccades of 7.5 degrees and saccades of 40 degrees is not just a quantitative difference (as they are assumed to be by Irwin (1998)) but a qualitative difference.

The second variable also has a potential problem, as the pronunciation of the non-words used in the experiment was arguably not controlled to a sufficient extent. It has been claimed that not all non-words are processed alike with respect to their pronunciation. A non-word such as *bint*, which has word neighbors that are inconsistent in their pronunciation (e.g.,



*pint, hint, mint*), takes longer to pronounce than one such as *tade* whose neighbors are consistent (Glushko, 1979, as cited in Rayner & Pollatsek, 1989). Rayner and Pollatsek (1989) argue that there are two systems active in pronouncing words: a direct lexical system, in which the pronunciation of words is looked up in the appropriate lexical entry, and a rule or analogy system, whereby the pronunciation of a word or non-word is generated by a system of rules or by a complex computation on a set of lexical and sublexical neighbors, or a combination of both. The results of Irwin's experiments might have been contaminated by the existence of such different types of non-words.

### 1.3.3. *Dependent variables.*

The dependent variable in Irwin's experiments was response time in the lexical decision task. There are some limitations that are inherent in the lexical decision task. First, the lexical decision task requires subjects to perform additional tasks such as preparing to press a button; it is possible that this may distort the course of the linguistic processing that would otherwise have taken place (Mitchell, 2004). Second, more importantly, it has long been noted that lexical decision latency inherently reflects not only the automatic lexical access but non-automatic processes such as response preparation and post-lexical process (e.g. Balota & Chumbley, 1984; Rayner & Pollatsek, 1989). For example, if there was a long sequence of non-words in an experimental session, the subject might then consciously wait for a word and make mental preparations to respond accordingly, thinking that a word is more likely after such a long sequence of non-words. And third, lexical decision tasks seem to be relatively prone to erratic responses. The subjects sometimes press the wrong button, show too long a latency (suggesting a possible lapse in attention), and make a decision before making a saccade. In fact, in Irwin's Experiment 1, 8.8% of the responses had to be discarded for some of these reasons.

Furthermore, Irwin's conclusion in this study was based on a null result; the prediction was that there should be no significant difference in response time between the long-saccade condition and the short-saccade condition if linguistic processing was not suppressed at all during saccades. As a result, it is not easy to evaluate the strength of his experimental evidence.

## 2. *A sketch of our experiment*

In this subsection, I propose a variant of Irwin's (1998) experiments that overcome their limitations; the goal here is to devise an experimental setup that gives us a way to study what happens when people read sentences as opposed to isolated words. The proposed experiment examines how the word immediately preceding a saccade location is processed during the saccade and immediately after the saccade.

### 2.1. *Materials.*

The stimuli used in the proposed experiment are pairs of sentences such as (1), which is an example adapted from Rayner and Duffy (1986).

- (1) The exhausted steward left the plane. / The exhausted student left the plane.

### 2.2. *Independent variables.*

What this experiment examines is, essentially, how the word immediately preceding a saccade location is processed during the saccade and immediately after the saccade. Therefore, the independent variables in this experiment will all concern the nature of the saccade and the nature of the word immediately preceding the saccade location. One independent variable is the length of the saccade, as in Irwin's (1998) experiments. Another independent variable is the corpus frequency of the word immediately preceding the saccade location. It has been noted in the literature that the corpus frequency of a word influences the length of the fixation on that word as well as the length of the fixation on the following words (e.g. Inhoff & Rayner, 1986; Rayner, 1998; Rayner & Duffy, 1986; Rayner, Fischer, & Pollatsek, 1998, Rayner, Sereno, & Raney, 1996, Vitu, 1991). For example, Rayner and Duffy (1986) reported that the fixation times were 30-90 ms longer on low-frequency words than on high-frequency words and the fixation times on a following word was 30-40 ms longer when it was preceded by a low-frequency word compared to when it was preceded by a high-frequency word. The reason why low-frequency words tend to prolong the fixation on the words following them is thought to be the following: when a reader processes a rarely used word, some cognitive processing concerning that word "spills over," that is, continues even after the eye has left the word. Such spill-over effects have long been observed in numerous other studies in eye movements as well (e.g. Balota, Pollatsek, & Rayner, 1985, Murray, 2000; Rayner, 1998).

Those stimulus words, the crucial pre-saccade words, should be carefully designed with respect to their length (possibly also syllable length) and (contextual) plausibility. Furthermore, the location of those pre-saccade stimulus words within the experimental sentences (or, equivalently, the location of the saccade in each experimental sentence) needs to be controlled; for instance, the amount of linguistic material preceding the stimulus word should be more or less constant across experimental conditions.

### 2.3. *Dependent variables.*

The dependent variable in this experiment will be the reading time after the presentation of the stimulus words, as measured by several eye tracking measures. Subjects' entire reading activity will be continuously recorded by an eye tracker and they do not have to make any extra decision during reading.

One problem that might arise in interpreting the result of such an experiment is the following. The relationship between eye movement and cognitive processing is now known not to be as tight as used to be assumed before (e.g. Just & Carpenter, 1980). Not all comprehension processes that operate on a word are initiated immediately after the fixation on that word; a reader processes a word not just while the eyes remain fixated on that word, because parafoveal attention goes beyond the fixation location and processing of a word could continue even after the eyes have left that word; and the control of the eye position is not always under cognitive control (Irwin, 2004; Vonk & Cozijn, 2003). Moreover, readers might employ voluntary, strategic eye movements, with the result that the location of their gaze will be distinct from the location that incurred processing difficulty (Liversedge, Paterson, & Pickering, 1998).

However, this problem is not an insurmountable problem. As the perceptual span is relatively small in the case of reading compared to scene processing, the fixation location provides a reasonably accurate estimate of the location of cognitive processing (Irwin, 2004). Furthermore, since the spill-over effect is thought to occur at a relatively early and automatic stage (Rayner & Duffy, 1986), it can reasonably be assumed that there will be relatively little effect of reader strategies.

If the process of reading comprehension is not suppressed during saccades, the reading time of the region after the saccade should be shorter when subjects have to execute a longer saccade, because the reader will have more time to process the region preceding the saccade location during longer saccades than during shorter saccades. Other predictions of the hypothesis will be discussed in the following section.

### 3. Method

This section provides a detailed description of our experiment.

#### 3.1. Subjects.

13 female and 14 male subjects participated in this experiment<sup>1</sup>. The subjects were either undergraduate or postgraduate students at the University of Edinburgh whose native language was English and had no background in psychology or linguistics. The age of the participants ranged from 18 to 33 (mean = 23.3, SD = 3.50). They had either uncorrected or corrected normal vision and were not aware of the purpose of the experiment. The subjects were paid 5 pounds for their participation.

#### 3.2. Materials.

66 pairs of stimuli were constructed. All of them are listed in the Appendix. The first half of each sentence, which was presented before the saccade location, consisted of 3 or 4 words (mean = 3.09, SD = 0.29) and 15 to 17 characters (mean = 16.45, SD = 0.58), while the second half of the sentence, which was presented after the saccade location, consisted of 2 to 4 words (mean = 2.80, SD = 0.50) and 14 to 17 characters (mean = 15.58, SD = 0.96). The vision angle per character in the sentence was approximately 0.31 degrees horizontally and 0.45 degrees vertically.

Each pair of sentences consisted of sentences that differed from each other only in one of the words used (e.g., *The large prison remained empty.* vs. *The large hangar remained empty.*). The words that differentiated the two sentences in a pair, which always immediately preceded the saccade location, will be called *target words* hereafter. In each pair of target words, one was always a high-frequency word (that is, a word that is used frequently), such as *prison*, and the other was always a low-frequency word (that is, a word that is used only infrequently), such as *hangar*; it was expected that the low-frequency target words would exert a stronger effect on the processing of the ensuing material than the high-frequency target words. The high-frequency and low-frequency words were matched in both the number of characters contained and the number of syllables contained.

The corpus frequency of each target word was determined using the Kucera-Francis frequency norms (Kucera & Francis, 1967) and the Brown verbal frequency ratings (Brown, 1984), obtained from the MRC database (Coltheart, 1981; Wilson, 1988). The frequency

---

<sup>1</sup> We removed from all of the analyses one female and two male subjects whose eye movements were not recorded on the sentence strings by the eye tracker at all in more than 25 % of the trials.

rating for each high-frequency target word was between 20 and 847 (mean = 127.41, SD = 133.36) in Kucera-Francis frequency ratings and between 0 and 533 (mean = 21.42, SD = 66.27) in Brown verbal frequency ratings, while the frequency rating for each low-frequency target word was between 0 and 2 (mean = 0.21, SD = 0.57) in Kucera-Francis frequency ratings and 0 in Brown verbal frequency ratings.

It was also verified that each high-frequency target word is a word that is acquired early and is familiar in its printed form and that each low-frequency target word is a word that is acquired relatively late and is relatively unfamiliar in its printed form. The age of acquisition of each word was determined by referencing the figures of Gilhooly and Logie (1980), multiplied by 100 in the MRC database (Coltheart, 1981; Wilson, 1988). The ratings for high-frequency target words were between 0 and 564 (mean = 283.44, SD = 164.94), while the ratings for low-frequency target words were out of the range of the database and always 0. The familiarity of each target word in its printed form was determined by merging the following three sets of familiarity norms: Paivio, Yuille and Madigan (1968), Toglia and Battig (1978) and Gilhooly and Logie (1980), obtained from the MRC database (Coltheart, 1981; Wilson, 1988). The ratings for high-frequency target words were between 0 and 634 (mean = 524.67, SD = 141.29), while the ratings for low-frequency target words were out of the range of the database and always 0. On the basis of these criteria, a word that was used in Rayner and Duffy (1986) as a low-frequency word was disqualified as a low-frequency target word, since the age of acquisition rating of that word was relatively low and the familiarity rating of that word was relatively high (508 and 291 respectively).

Target words were always at least 6 characters long, so that they would not be likely to be skipped over. It has long been established that there is a systematic relationship between the length of a word and the probability with which it is skipped. Short words (3 characters or less) are much more likely to be skipped than longer words (Rayner & Sereno, 1994). Words 6 characters long are almost always fixated and words 8 characters or longer are rarely if ever skipped (Rayner & McConkie, 1976). The target words were therefore between 6 and 8 characters (mean = 6.59, SD = 0.72) long, and contained either 2 or 3 syllables (mean = 2.17, SD = 0.38). As already noted above, target words that were paired always had the same number of characters and the same number of syllables; for example, *prison* and *hangar* both contain 6 characters and 2 syllables.

Likewise, the word immediately after the saccade, which will be called a *spill-over word* hereafter, was always more than 4 characters long (min. = 4, max. = 10, mean = 5.83, SD = 1.56), so that it would not easily be skipped.

The first half of each sentence fell into one of the following groups.

- (1) 31 pairs of word sequences consisting of a non-quantificational determiner, an adjective, and a noun (e.g., *The young cousin*). This is a type of string that is also used in the

previous experiments concerned with the spill-over effect (Schilling, Rayner & Chumbley, 1998; Rayner & Duffy, 1986).

- (2) 13 pairs of noun phrases preceded by an adverb (e.g., *Soon the audience*).
- (3) 10 pairs of partial verbal phrases consisting of (i) a verb that obligatorily subcategorizes for an NP and a PP and (ii) an NP. One example of this type of string is the following: *I put the picture*. (The PP subcategorized by the verb appeared after the saccade location.) The subcategorization frame of a verb was determined by inspecting the electrical COMLEX syntax lexicon (Grishman, Macleod & Meyers, 1994), and it was ensured that the sequence made up of a verb and the following NP cannot be construed as a complete VP.
- (4) 7 pairs of noun phrases beginning with a quantifier (e.g., *All those answers*)
- (5) 3 pairs of conjoined noun phrases (e.g., *Maths and history*).
- (6) 2 pairs of gerundives (e.g., *Finding a market*).

The unpredictability of the target words was checked by Cloze tests, in which 23 participants, who were different from the participants of the main experiment, tried to guess the target words after having read the portion of the sentences that preceded them. The target words were never correctly guessed by more than 19 percent of the participants.

### 3.3. Assignment of the materials.

Four lists<sup>2</sup> of experimental sentences were created using the Latin Square design, so that the following conditions would be satisfied: (i) for each sentence pair, one fourth of the subjects read the sentence with a high-frequency target word making a long saccade, one fourth of the subjects read the same sentence making a short saccade, one fourth of the subjects read the sentence with a low-frequency target word making a long saccade, and one fourth of the subjects read the same sentence making a short saccade, (ii) each fourth of the subjects experiences equal numbers of each of the four experimental conditions (i.e. high-frequency vs. low-frequency and long saccade vs. short saccade), (iii) each subject sees any given sentence frame (i.e. a sentence minus the target word) once and only once, and (iv) each subject sees any given target word at most once.

As noted above, the number of experimental sentence pairs was 66, not a multiple of 4. This created a slight imbalance in the experimental design, but all the measurements generated were used in the final analysis. To be more specific, (i) one fourth of the subjects read 17 sentences with a high-frequency target word making a long saccade, 17 sentences

---

<sup>2</sup> Since we removed from all of the analyses one female and two male subjects as was explained in the subject section, the result of 12 female and 12 male subjects remained to be analyzed. Within those remaining subjects, 3 female subjects and 3 male subjects had been assigned to each list.

with a high-frequency target word making a short saccade, 16 sentences with a low-frequency target word making a long saccade, and 16 sentences with a low-frequency target word making a short saccade, (ii) one fourth of the subjects read 16 sentences with a high-frequency target word making a long saccade, 17 sentences with a high-frequency target word making a short saccade, 17 sentences with a low-frequency target word making a long saccade, and 16 sentences with a low-frequency target word making a short saccade, (iii) one fourth of the subjects read 16 sentences with a high-frequency target word making a long saccade, 16 sentences with a high-frequency target word making a short saccade, 17 sentences with a low-frequency target word making a long saccade, and 17 sentences with a low-frequency target word making a short saccade, (iv) one fourth of the subjects read 17 sentences with a high-frequency target word making a long saccade, 16 sentences with a high-frequency target word making a short saccade, 16 sentences with a low-frequency target word making a long saccade, and 17 sentences with a low-frequency target word making a short saccade.

Four sentences which embody the four experimental conditions were used for a practice trial, prior to the experiment. None of the four sentences were used in the main experiment.

The order of presentation of the sentences in each list was randomized on a subject-by-subject basis. The data collection for each subject was accomplished in a single 40-minute-long session.

### 3.4. *Apparatus.*

A Dell Dimension DIM 3000 (Intel Pentium ® 4 CPUs 2.80 GHz equipped with 512MB of RAM) microcomputer controlled a head mounted eye tracker, EyeLink® II (manufactured by SR Research Ltd.), and recorded outputs from it. Subjects responded to comprehension questions by pressing an HID compliant hand-held game-pad switch. Stimuli were presented at a refresh rate of 120 Hz on an Iiyama HM204D/DT colour monitor. A Dell Precision PWS 370 (Intel Pentium ® 4 CPUs 3.20 GHz equipped with 1GB of RAM and an NVIDIA Quadro FX3400 graphics adaptor) controlled stimulus presentation.

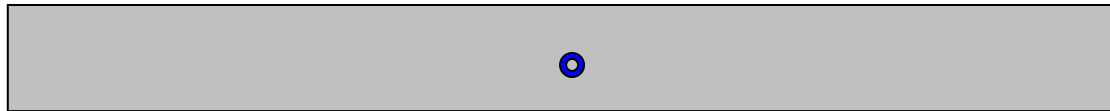
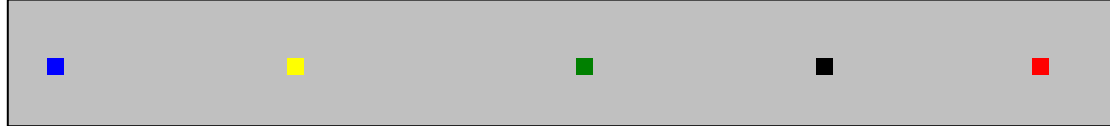
The eye tracker was mounted on eye camera frames that were held in place on the subject's head by a headband. The eye tracker was configured to record movements of the subjects' dominant eye only. Eye position was sampled 500 times per second. A chin rest was used to keep the subject's head steady during the experiment. Subjects completed 13 point grid calibration and validation sequence twice, first before the experiment started and then after a break in the middle of the experiment. The accuracy of the eye tracker was  $\pm 0.5$  deg.

During the experiment, subjects were seated 42.5 cm from the display monitor. At this viewing distance, the total display area subtended 50.95 deg of visual angle horizontally and 39.36 deg vertically. Each letter string subtended approximately 0.31 deg horizontally and 0.45 deg vertically. A box whose dimensions were 0.40 deg horizontally by 0.51 deg vertically and was almost the same size with one letter was used as a target box and also as a fixation box. Calibration points used during the experiment were concentric circles whose outer circle subtended 16 pixels and whose inner circle subtended 4 pixels (approximately 0.8 deg and 0.2 deg respectively). The stimuli were presented on a screen which consisted of 1024 pixels horizontally and 768 pixels vertically, using the Courier New font in 8 point size and normal font style. The display background was light grey (Red, Green and Blue (here henceforth RGB) = 153,153,153; Hue, Saturation and Brightness (here henceforth HSB) = 0, 0, 60), while the letters and calibration points were black (RGB = 0, 0, 0; HSB = 0, 0, 0) and the target boxes were dark blue (RGB = 0, 0, 255; HSB = 239, 100, 100). The fixation boxes were five coloured; dark blue (RGB = 0, 0, 255; HSB = 239, 100, 100), yellow (RGB = 255, 255, 0; HSB = 60, 100, 100), green (RGB = 0, 153, 0; HSB = 120, 100, 60), black (RGB = 51, 51, 51; HSB = 0, 0, 20) and red (RGB = 255, 0, 51; HSB = 347, 100, 100).

### 3.5. *Procedure.*

Each experiment was started by determining the subject's dominant eye. The dominant eye can be defined as the eye that looks directly at an object while the non-dominant eye can be defined the eye that looks at the same object at a slight angle. This small difference between the two eyes provides depth perception, the ability to determine the distance between objects. Conventionally, the movement of the dominant eye alone is recorded, when only the movement of one eye needs to be recorded. This convention was followed in the present experiment. A test for eye dominance was conducted according to the guidelines of the Council for Refractive Surgery Quality Assurance, which is an independent, non-profit patient/consumer public benefit health organization (according to the [USAeyes.org](http://USAeyes.org) website). Participants, with both their eyes open, raised their index finger and pointed to an object in the distance. They continued to look at and point at the object and cover the left eye. They checked if their index finger seemed to move off the target with their left eye covered. They did the same with the right eye. Then they reported the dominant eye which was the eye they were using when the index finger did not seem to move.



**Drift Correction:****Fixations:****Fixation (1500ms):****First half of sentence and saccade target:****Saccade and second half of sentence:****Comprehension question:****Fig.1.** Schematic illustration of the main procedure in the experiment.

The main procedure is shown in Fig. 1. Subjects were seated 42.5 cm from the display monitor and began each trial by fixating a circle in the middle of the screen, so that a drift correction could be performed. Then five coloured square boxes appeared which were separated by 12 deg of visual angle on the display. The subjects fixated some of these coloured square boxes which were mentioned by the experimenter. Following the trial of fixation using the five fixation squares, a single fixation square appeared on the left side of the display which indicated the starting point of the sentence. The subject fixated that square, or the area roughly 1.7 degrees wide and 1.5 degrees high which surrounded that square,<sup>3</sup> for 1500 ms. Then the first half of the sentence appeared on the left side of the display. A saccade target square was presented on the right side of the display either 10 deg or 40 deg

<sup>3</sup> In other words, this larger square area, which was implicit and not indicated on the screen, was treated as the fixation box. This procedure was adopted because the pilot experiment had shown that some people tend not to fixate the inside of the fixation box but look slightly away from the box.

away from the end of the first half of the sentence at the same time the first half of the sentence was presented. The subject was instructed to move his eyes onto the target square as soon as he finished reading the first half of the sentence silently. The first half of the sentence was erased and the second half of the sentence was displayed when the subjects' eyes overrode the invisible boundary which was located at the end of the first half of the sentence. The left edge of the second half was located where the saccade target square had been shown. Subjects then read the second half of the sentence silently, after making a saccade. They were instructed to read as quickly but as accurately as possible. The subjects were instructed to press any key on the hand-held game-pad switch as soon as he finished reading the second half of the sentence. The subjects were asked a simple yes/no comprehension question at this time on 33% of the trials; the subjects were expected to press either the yes or the no button on the game-pad switch to answer the questions.

Each subject completed one block of 66 experimental trials. Subjects completed one practice block with 4 trials (one high-frequency target word followed by a long saccade, one high-frequency target word followed by a short saccade, one low-frequency target word followed by a long saccade and one low-frequency target word followed by a short saccade, sequenced randomly) before beginning the experimental trials.

### 3.6. Data Analyses.

Of primary interest is whether low-frequency target words do affect the processing of the following words to a greater extent than high-frequency target words do and whether the magnitude of such spill-over effects is affected by saccade distance.

To investigate these issues, we analyzed the following five kinds of eye-tracking measures. The first four kinds of measurements were taken in order to analyze how the target words themselves were processed. The fifth kind of measurement was taken in order to analyze how the spill-over words, that is, the words immediately following the saccade location, were processed. In what follows, the region of the screen covered by a target word or a spill-over word will be referred to as a target-word region and a spill-over word region, respectively.

We analyzed the following three eye tracking measures for the target word traditionally used in other literature (e.g. Rayner & Duffy, 1986; Murray, 2000): *first fixation duration*, which is the length of the first fixation in the target-word region, *gaze duration*, which is the sum of the lengths of all fixations within the target-word region starting with the subject's first fixation within the region until the subject's gaze leaves the region either to the left or right for the first time, and *total time*, which is the sum of the lengths of all fixations within

the target-word region starting with the subject's first fixation within the region until the subject's gaze leaves the region either to the left or right for the last time.

Trials like the following were discarded and not used in the analysis:

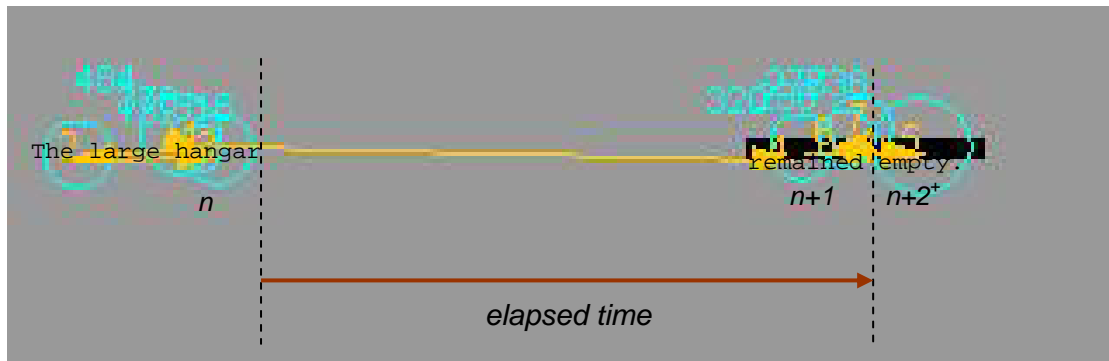
- (1) trials in which the subject's eye position was not recorded at all and
- (2) trials in which the subject's eyes did not land on the target word.

In total, the mean of the percentage of acceptable trials in each condition of each subject was 92.4% (SD = 13).

If a fixation was less than 60 ms long and within one character space of the previous or next fixation, it was assimilated into the largest fixation within one character space. All remaining fixations shorter than 60 ms were disregarded; they were not included in the calculation of any type of fixation duration, and their existence was ignored for the purpose of distinguishing successful trials (i.e. trials to be included in the analysis) from unsuccessful ones (i.e. trials to be discarded). The decision to disregard fixations shorter than 60 ms was based on the fact that readers are able to extract (some) information only during fixations longer than 60 ms (Liversedge et al., 2004; Rayner, Liversedge, White, & Vergilino-Perez, 2003). In addition, fixations longer than 1000 ms were regarded as outliers and excluded from analysis.

The fourth measurement to be analyzed in the experiment was the pupil size as it was recorded when the subject fixated the target word for the first time. It is known that the difficulty of the subject's task could be reflected in the size of the pupil (e.g., Takahashi, Nakayama, & Shimizu, 2001).

The fifth measurement to be analyzed in the experiment, which will be referred to as *elapsed time* in what follows, is the time elapsed from the beginning of the saccade launched directly from the target word till the start of the first fixation on the word immediately following the spill-over word. In other words, the elapsed time is the duration of the saccade plus the duration of the first fixation on the spill-over word. Fixations on the spill-over word after the first fixation, if there were any, were not included in the elapsed time; this is because the processing of the target word is not expected to continue so long after the eyes have left the target word (Rayner & Duffy, 1986).



**Fig.2.** An example of the measurement of the elapsed time

In the example shown in Fig.2, the target word is *hangar*. We measured the time elapsed from the beginning of the saccade launched from this target word till the start of a fixation on the word *empty*, which immediately follows the spill-over word, *remained*.

Elapsed time was not measured for trials like (1) and (2) above, and it was also not measured for trials like the following:

- (3) trials in which the eyes made a regression onto a previous word (such as *The* or *large* in Fig. 2.) after landing on the target word, and then made a saccade directly from one of those previous words, without landing on the target word again and
- (4) trials in which the eyes did not land before or on the spill-over word and landed directly onto a following word (such as *empty* in Fig. 2) and never came back to fixate the spill-over word.

With these more stringent criteria, the mean of the percentage of acceptable trials in each condition of each subject was 83.3% (SD = 24).

We did not exclude from analysis those trials in which the eyes either over- or undershot the spill-over word, made a corrective movement to reach the spill-over word, and then moved further to the right; all fixations in such trials were dealt with the same way that fixations in other, more canonical trials were. However, in analyzing trials in which the eyes made a corrective movement to reach the spill-over word after overshooting and never moved further to the right, we excluded the last fixation on the spill-over word from analysis, since the length of such a fixation is expected to reflect what is called the *wrap up effect*, that is, readers' tendency to pause longer on a word that ends a clause or a sentence (Just & Carpenter, 1980; Hill & Murray, 2000; Rayner, Kambe, & Duffy, 2000; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989); during such a prolonged fixation that happens when a reader is about to finish reading a clause or a sentence, the reader is presumably processing not just the last word that has been looked at but also the meaning of the entire clause or the entire sentence that has been read.

#### 4. Results

Preliminary analysis of the eye-movement data showed that the mean distance of the initial saccade in the 10 deg saccade condition was 10.9 deg ( $SE = 0.1$  deg), whereas it was 39.1 deg ( $SE = 0.1$  deg) in the 40 deg saccade condition, which is 1.4 times<sup>4</sup> and 1.1 times longer than that of Irwin's (1998) Experiment 1, respectively. Mean saccade duration was 54.5 ms ( $SE = 0.6$  ms) in the 10 deg saccade condition and 126.7 ms ( $SE = 1.1$  ms) in the 40 deg saccade condition, which is 1.9 times and 1.3 times longer than that of Irwin's (1998) Experiment 1, respectively.

Analysis of the distribution of trials that had to be discarded (i.e. trials that fell into one of the four categories (1) – (4) defined in the previous section) showed that there was not a significant association between the acceptance rate of trials and the saccade length ( $\chi^2 = 1.02$ ,  $p > 0.2$ ), while there was a significant association between the acceptance rate of trials and the corpus frequency of the target words ( $\chi^2 = 8.38$ ,  $p < 0.01$ ). The association between the acceptance rate of trials and the corpus frequency of the target words was significant because high-frequency target words were more likely to be skipped.

What was of primary interest in this experiment was whether low-frequency target words affected the processing of the following words to a greater extent than high-frequency target words did and whether the magnitude of such spill-over effects was affected by saccade distance.

In the remainder of this section, ancillary results, namely results pertaining to the way the subjects looked at the target words themselves, will be presented first, and then primary results, namely results pertaining to the spill-over effects, will be presented.

##### 4.1. Ancillary results: how the subjects looked at the target words themselves.

The first fixation duration, gaze duration and total time were all expected to be longer for low-frequency target words than for high-frequency target words, and they were not expected to be influenced by saccade distance. Likewise, the pupil size was expected to be larger for low-frequency target words than for high-frequency target words.

---

<sup>4</sup> The saccade length in the short-saccade condition in Irwin (1998) was 7.5 degrees. All else being equal, therefore, the saccade duration in the short-saccade condition in the present experiment was expected to be about 1.33 times longer than the saccade duration in the short-saccade condition in Irwin's experiments.

**Table 1****Mean Reading Time (in ms) and Pupil Size as a Function of Saccade Length and Frequency of Target Word**

Eye tracking measures	Target word frequency	
	Saccade Length	High Low
First fixation duration		
Long	208 (5)	234 (6)
Short	213 (5)	232 (5)
Gaze duration		
Long	245 (7)	320 (9)
Short	255 (8)	332 (9)
Total times		
Long	250 (7)	336 (10)
Short	264 (8)	356 (10)
Pupil sizes at first fixation		
Long	1356 (28)	1356 (27)
Short	1363 (28)	1362 (28)

Table 1 presents the mean reading times and pupil size<sup>5</sup> of accepted trials as a function of the saccade length and the corpus frequency of the target word. The mean reading times are reported in milliseconds, and the figures in parentheses represent the corresponding standard errors. For each type of measurement, we conducted two 2-way repeated-measures Analyses of Variance (ANOVAs), one with subjects ( $F_1$ ) and one with items ( $F_2$ ) as the random factor, and both with the saccade length and the corpus frequency of the target word as the fixed factors. All the ANOVAs reported in this subsection except those pertaining to the pupil sizes were carried out on the log-transformed values of the measurements, as D'Agostino-Pearson tests indicated that the raw measurements in each experimental condition often deviated significantly from the normal distribution and that the log-transformed values mostly did not. On the other hand, the ANOVAs pertaining to the pupil sizes were carried out on the square roots of the measurements, as D'Agostino-Pearson tests indicated that the distribution of those square roots was closer to the normal distribution than that of the raw measurements in each experimental condition.

<sup>5</sup> The figures for pupil sizes are unitless numbers generated by EyeLink which represent the relative sizes of the measured lengths.

There was a significant main effect of target word frequency in all three types of reading times (first fixation duration:  $F_1(1, 23) = 29.11, p < 0.01$ ;  $F_2(1, 65) = 19.15, p < 0.01$ ; gaze duration:  $F_1(1, 23) = 76.13, p < 0.01$ ;  $F_2(1, 65) = 98.89, p < 0.01$ ; and total time:  $F_1(1, 23) = 78.27, p < 0.01$ ;  $F_2(1, 65) = 103.59, p < 0.01$ ); the subjects looked at low-frequency target words longer than they looked at high-frequency target words. The main effect of saccade length was not significant in the case of first fixation duration ( $F_1(1, 23) = 1.45, p > 0.2$ ;  $F_2(1, 65) = 1.02, p > 0.3$ ), but significant in analysis by subjects in the case of gaze duration ( $F_1(1, 23) = 5.17, p < 0.05$ ;  $F_2(1, 65) = 3.94, p = 0.051$ ) and significant in both analysis by subjects and analysis by items in the case of total time ( $F_1(1, 23) = 8.61, p < 0.01$ ;  $F_2(1, 65) = 6.60, p < 0.02$ ); gaze duration and total time were both longer in the short-saccade condition than in the long-saccade condition. There was no interaction between target word frequency and saccade length in any of the three types of measurements (first fixation duration:  $F_1(1, 23) = 1.17, p > 0.2$ ;  $F_2(1, 65) = 0.98, p > 0.3$ ; gaze duration:  $F_1(1, 23) = 0.17, p > 0.6$ ;  $F_2(1, 65) = 0.22, p > 0.6$ ; total time:  $F_1(1, 23) = 0.02, p > 0.8$ ;  $F_2(1, 65) = 0.02, p > 0.8$ ).

There was a non-significant main effect of target word frequency in pupil sizes when the subjects looked at the target word at first glance ( $F_1(1, 23) = 3.43, p > 0.07$ ;  $F_2(1, 65) = 0.01, p > 0.9$ ) and no main effect of saccade length ( $F_1(1, 23) = 0.01, p > 0.9$ ;  $F_2(1, 65) = 0.03, p > 0.8$ ). There was no interaction between target word frequency and saccade length ( $F_1(1, 23) = 0.87, p > 0.3$ ;  $F_2(1, 65) = 0.11, p > 0.7$ ).

The fact that the subjects looked at the low-frequency target words longer than they looked at the high-frequency target words was an expected one, and shows that the corpus frequency of the target words was successfully controlled. On the other hand, the fact that subjects fixated the target words longer in the short saccade condition was not an expected one<sup>6</sup>.

#### 4.2. Primary results: Spill-over effects after the saccade.

The elapsed time is expected to be longer when the target word is a low-frequency word, due to spill-over effects, than when the target word is a high-frequency word. The elapsed time is also expected to be longer when the saccade length is long, because a saccade over a longer distance takes longer to make than a saccade over a shorter distance.

The key prediction of the experiment is concerned with whether an interaction between the length of the saccade and the corpus frequency of the target word is to be detected by the

---

<sup>6</sup> Closer examination reveals that, while the gaze duration and total time were longer in the short saccade condition, the first fixation duration was roughly the same between the short saccade condition and the long saccade condition. I have to leave the explanation of these observations to future work.

ANOVA whose independent variables are the length of the saccade and the corpus frequency of the target word and whose dependent variable is the elapsed time.

If cognitive processing is completely suppressed during saccades, then there should be no interaction between the length of the saccade and the corpus frequency of the target word; a low-frequency word is expected to take just as much longer than a high-frequency word in the long-saccade condition as in the short-saccade condition.

On the other hand, if processing is not completely suppressed during saccades, then there should be some interaction between the corpus frequency of the target word and the length of the saccade. More specifically, the difference in elapsed time between the high-frequency word and the low-frequency word is expected to be smaller in the long-saccade case, compared to the short-saccade case. To see why this should be the case, consider an extreme case, in which the saccade is sufficiently long, so that the processing of the target word is completed during the saccade, no matter whether the target word is a high-frequency word or a low-frequency word. In a situation involving such an extremely long saccade (during which the reader continues linguistic processing, by assumption), there will not be any difference in elapsed time between high-frequency target words and low-frequency target words (whereas there is expected to be a difference in elapsed time between the two types of target words in the short-saccade condition).

**Table 2**

**Mean Elapsed Time (in ms) and Percentage Correct as a Function of Saccade Length and Corpus Frequency of Target Word**

Eye tracking measures		Target Word Frequency		
		High		Low
Saccade Length	ET	% correct	ET	% correct
Elapsed times				
Long	505 (12)	96.8	547 (13)	92.8
Short	373 (9)	95.2	449 (13)	96.4

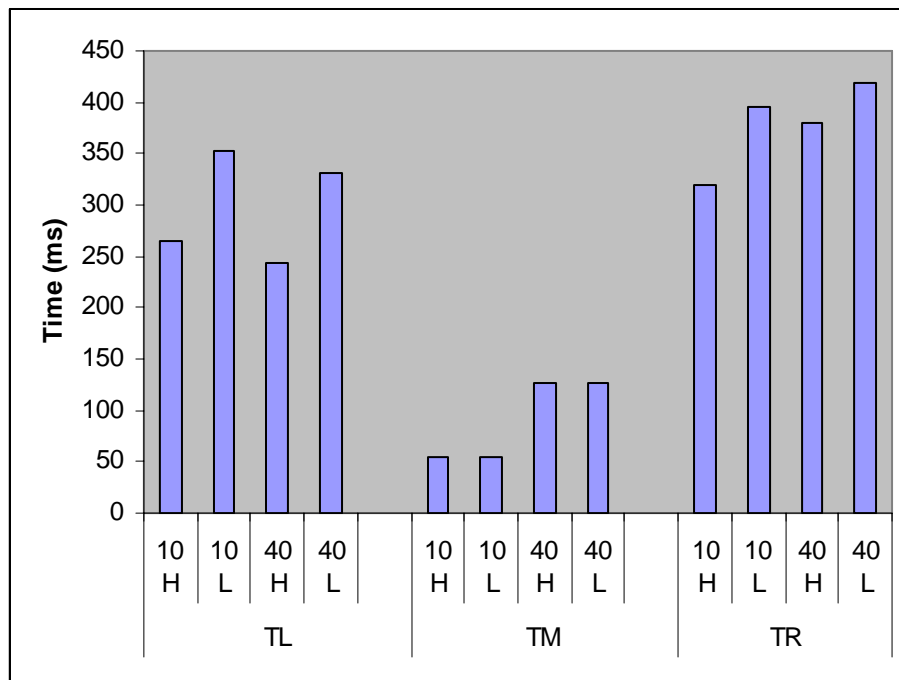
Table 2 presents the mean elapsed time (ET) of the accepted trials and the percentage of the correct responses to the comprehension questions as a function of the saccade length and the corpus frequency of the target word. The averages of elapsed time are reported in milliseconds, and the corresponding standard errors are shown in parentheses. For each condition, we conducted two 2-way repeated-measures ANOVAs, one with subjects ( $F_1$ ) and one with items ( $F_2$ ) as the random factor and both with the saccade length and the corpus



frequency of the target word as fixed factors. All the ANOVAs reported in this subsection, except for the ANOVA concerning post-saccadic processing time (TR), were carried out on the log-transformed values of the measurements, as D'Agostino-Pearson tests indicated that the raw measurements in each experimental condition often deviated significantly from the normal distribution and that the log-transformed values mostly did not. The ANOVA concerning post-saccadic processing time (TR), on the other hand, was carried out on the raw measurements, because they were found not to deviate from the normal distribution any more than the log-transformed values.

There was a significant main effect of target word frequency ( $F_1(1, 23) = 21.53$ ,  $p < 0.01$ ;  $F_2(1, 65) = 26.01$ ,  $p < 0.01$ ); as predicted, the elapsed time was longer when the target word was a low-frequency word. There was also a significant main effect of saccade length ( $F_1(1, 23) = 19.45$ ,  $p < 0.01$ ;  $F_2(1, 65) = 110.14$ ,  $p < 0.01$ ); the elapsed time was significantly longer in the 40-deg saccade condition than in the 10-deg saccade condition. Furthermore, there was an interaction between target word frequency and saccade length ( $F_1(1, 23) = 4.72$ ,  $p < 0.05$ ;  $F_2(1, 65) = 3.10$ ,  $p = 0.081$ ), although this was significant only in  $F_1$ .

Roughly emulating Irwin (1998), the elapsed time was broken down into two component measures (see also Sanders, 1963, 1970, as cited in Irwin, 1998): TM, *time moving*, the duration of the saccade; and TR, *time right*, the time that elapsed between the subject's eye landing on or near the spill-over word and the start of the fixation of the word immediately after the spill-over word. By definition, for each trial, the value of the measurement that we have decided to refer to as elapsed time was equal to the sum of the TM and the TR.



**Fig.3.** Mean total time the subjects fixated the target word (TL), mean saccade duration (TM), and mean post-saccadic processing time (TR) as a function of saccade length and corpus frequency of target word

The result of this breakdown is shown in Fig.3, as a function of saccade length and the frequency of the target word. In Fig.3, each column is tagged with the combination of 10 or 40 and H or L. 10 and 40 indicate the saccade length and refer to 10 degrees and 40 degrees of visual angle respectively. H and L refer to the high corpus frequency and the low corpus frequency of the target word respectively. What is called TL (again emulating Irwin (1998)) in this figure is the mean total time<sup>7</sup> the subjects fixated the target word in each experimental condition.

Separate 2-way repeated-measures ANOVAs were conducted on the TL, TM and TR (from acceptable trials only), with the saccade length and the corpus frequency of the target word as the independent variables. As indicated above, these ANOVAs were carried out on the log-transformed values, with the exception of the ANOVA pertaining to TR, which was carried out on the raw figures.

TL, the time spent fixating the target word before the saccade onset, showed a significant main effect of target word frequency ( $F_1(1, 23) = 86.81, p < 0.01$ ;  $F_2(1, 65) = 95.54, p < 0.01$ ) and a significant main effect of saccade length ( $F_1(1, 23) = 6.09, p < 0.05$ ;

<sup>7</sup> Since trials which fall into categories (3) and (4) defined in 3.6 are excluded here, the figures are slightly different from those shown in 4.1.

$F_2(1, 65) = 10.37, p < 0.01$ ). There was no interaction between target word frequency and saccade length ( $F_1(1, 23) = 0.03, p > 0.8$ ;  $F_2(1, 65) = 0.13, p > 0.7$ ).

The mean saccade duration (TM) showed a marginally significant main effect of target word frequency ( $F_1(1, 23) = 3.24, 0.09 < p < 0.1$ ;  $F_2(1, 65) = 4.09, p < 0.05$ ), but the differences among conditions were so small (less than 2 ms) that it is difficult to assign much psychological significance to this. TM showed a significant main effect of saccade length ( $F_1(1, 23) = 1550.32, p < 0.01$ ;  $F_2(1, 65) = 5468.21, p < 0.01$ ). There was no interaction between frequency of target word and saccade length conditions ( $F_1(1, 23) = 2.31, p > 0.1$ ;  $F_2(1, 65) = 0.95, p > 0.1$ ).

The mean of post-saccadic processing time (TR) showed a significant main effect of the corpus frequency of the target word ( $F_1(1, 23) = 17.9, p < 0.01$ ;  $F_2(1, 65) = 31.0, p < 0.01$ ). There was a marginally significant main effect of the saccade length ( $F_1(1, 23) = 2.4, p > 0.1$ ;  $F_2(1, 65) = 13.7, p > 0.7$ ). The interaction between target word frequency and saccade length was also marginally significant ( $F_1(1, 23) = 4.1, 0.05 < p < 0.06$ ;  $F_2(1, 65) = 3.1, p = 0.081$ ).

As noted in the previous section, we did not exclude from analysis those trials in which the eyes either over- or undershot the spill-over word, made a corrective movement to reach the spill-over word, and then moved further to the right. There was a significant association between the saccade length and whether the eyes landed precisely on the spill-over word without undershooting or overshooting ( $\chi^2 = 191.49, p < 0.001$ ). The subjects' eyes overshoot or undershot more often in the long-saccade condition than in the short-saccade condition, presumably prolonging the reading time associated with the spill-over word region. This association between the accuracy of a saccade and its length is not unexpected. For one thing, a longer saccade is presumably harder to execute than a shorter saccade and the accuracy of the former is thus naturally expected to be lower than that of the latter. Furthermore, several previous studies, using non-linguistic target stimuli, indicate that the accuracy of a saccade is affected by which portion of the computer screen the saccade target is located (e.g. Kapoula & Robinson, 1986); centripetal saccades that return the eyes to the center of the screen tend to be more accurate than centrifugal saccades that move the eyes away from the center of the screen. The short-saccade condition in our experiment probably involved centripetal saccades and the long-saccade condition probably involved centrifugal saccades in this sense.

On the other hand, the association between the corpus frequency of the target word and whether or not the eyes either undershot or overshoot was not significant ( $\chi^2 = 3.10, p > 0.08$ .) Thus the existence of undershooting and overshooting arguably has not seriously affected the interaction result presented above. (This last point will be elaborated on in the next section.)

## 5. Discussion

In this section, I will attempt to interpret the results of the present experiment, and will also discuss the possible future directions of this line of study.

### 5.1. Interpretations of results.

Analysis of the first fixation duration, gaze duration and total time for the target word showed that subjects looked at infrequently used target words longer than frequently used target words, a result consistent with other eye movement studies on lexical frequency (Rayner & Duffy, 1986; Schilling et al., 1998). Moreover, the processing time for the words after saccades (the spill-over words) was longer when the subjects had read an infrequently used target word than when they had read a frequently used target word, a result indicating that processing of the target words spilled over onto the words after saccades at least in the case of low-frequency target words, as was expected. These results indicate, as was already noted in the last section, that the way the corpus frequency of the target words was controlled in the present experiment was adequate; the words that were supposed to be high-frequency words behaved as high-frequency words are expected to and the words that were supposed to be low-frequency words behaved as low-frequency words are expected to.

There was an interaction effect between saccade length and target word frequency on elapsed time (the time elapsed from the beginning of the saccade launched directly from the target word till the start of the first fixation on the word immediately following the spill-over word), although this effect was significant only in  $F_1$ . Specifically, the difference in elapsed time between the high-frequency target word condition and the low-frequency target word condition was smaller in the 40-deg saccade condition than in the 10-deg saccade condition. This is consistent with what we would expect if cognitive processing for reading was not completely suppressed during saccades. Let me recapitulate why this is so, in somewhat more detail than in the previous section. If cognitive processing for reading were completely suppressed during saccades, the length of a saccade would not have any effect on the difference in elapsed time between the low-frequency target word condition and the high-frequency target word condition; the only effect that a saccade could have on elapsed time would be to add a fixed amount of time (i.e. either the time necessary to complete a 10-deg saccade or the time necessary to complete a 40-deg saccade) to the elapsed time, in any of the experimental conditions. If, on the other hand, cognitive processing for reading is not suppressed during saccades, processing of high-frequency target words, which is easier, is expected to be completed during the saccade more often than processing of low-frequency target words, which is harder and generally takes longer. When processing of the target word

is completed during the saccade, the processor has to “wait”, with nothing to do, until the eyes land on the spill-over word. Therefore, in the high-frequency target word condition, the elapsed time is expected to be artificially prolonged in the long saccade condition. The same artificial prolongation might occur in other conditions as well, but it is not expected to occur to the same extent as in the case involving a high-frequency target word and a long saccade. Since high-frequency words tend to be processed faster in general, this artificial prolongation leads to a situation in which the difference between high-frequency target words and low-frequency target words is smaller in the long-saccade condition, where more artificial prolongation is expected to take place, than in the short-saccade condition, where less artificial prolongation is expected to take place.

Can we conclude, then, that linguistic processing was not completely suppressed during saccades when people read sentences, just as it was not suppressed in Irwin’s (1998) experiments involving isolated words? There are in fact several complications that need to be discussed before we could reach that conclusion.

First, contrary to the result of the current experiment, the experiments of Irwin (1998) did not find an interaction effect between word type and saccade length on the course of linguistic processing. This fact might appear to conflict with the interpretation offered above. Closer examination reveals, however, that there is no contradiction. In Irwin’s experiments in question, there was arguably no reliable main effect of word type on the course of linguistic processing in the first place. In Irwin (1998), it is reported that the difference in TL between words and non-words was significant in Experiment 1 but not in Experiment 2, in which Irwin attempted to reduce the time for saccade programming (which is included in TL) by showing the target box earlier than the letter string. The difference in TR between words and non-words was likewise significant in Experiment 1 but not significant in Experiment 2. The difference in total response time between words and non-words was not significant in Experiment 1 or Experiment 2. These observations indicate that, in Irwin’s experiments in question, there was no reliable difference between words and non-words as far as their effect on the time course of linguistic processing is concerned. This probably means one of the following two things: either the difference between words and non-words is quite different from the difference between high-frequency words and low-frequency words, or the choice of experimental non-words was not entirely adequate in his experiment, as was noted as a possibility in the section concerned with the potential limitations of Irwin’s (1998) method. No matter which of these two interpretations turns out to be correct, there is nothing that contradicts the above interpretation of the present experiment.

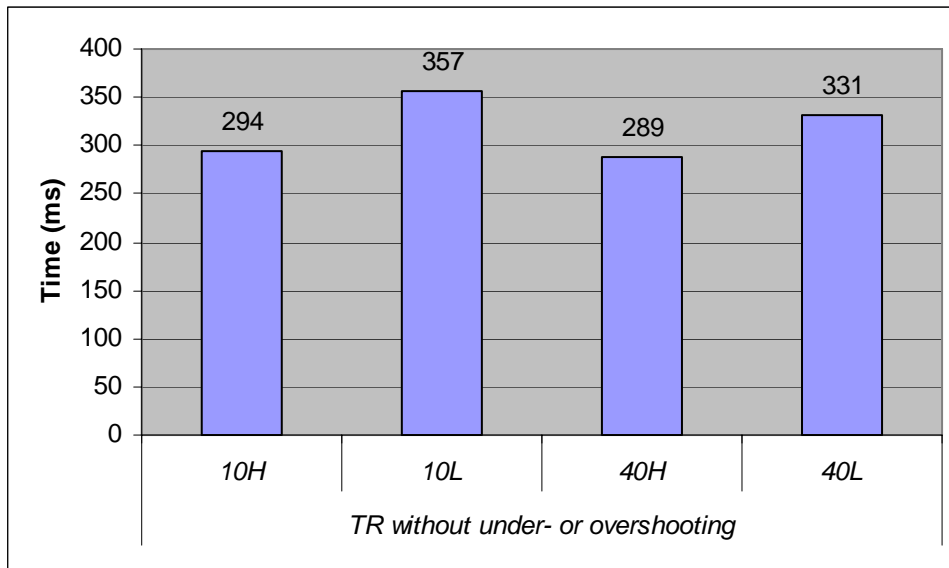
The second complication to be discussed is the following. As shown in Fig. 3 of the last section, TR (which is elapsed time minus saccade duration, in other words, the time that elapsed between the subject’s eye landing on or near the spill-over word and the start of the

fixation of the word immediately after the spill-over word) was on average longer in the 40-deg condition compared to the 10-deg condition. This is the opposite of what we would expect on the assumption that linguistic processing was not suppressed during saccades; TR should be shorter in the long-saccade condition, because the processor should be able to do more processing during the saccade in the long-saccade condition than in the short-saccade condition. It is also not what we would expect on the assumption that linguistic processing *was* suppressed during saccades; on that assumption, there should be no difference in TR between the long-saccade condition and the short-saccade condition. However, this observation seems to be more problematic to the hypothesis that linguistic processing is not suppressed during saccades than to the opposite hypothesis, and thus calls for a close examination.

The situation is further complicated by the fact that, in Irwin's (1998) experiments, the mean TR was shorter in the long-saccade condition than it was in the short-saccade condition. This adds to the impression that TR should have been shorter in the long-saccade condition in our experiment as well.

This complication, however, does not undermine the conclusion that linguistic processing was not suppressed during saccades. The reason that TR was longer in the long-saccade condition could simply be that undershooting and overshooting of the eyes took place more often in the long-saccade condition compared to the short saccade condition; undershooting and overshooting have the effect of making TR longer since the position of the eyes has to be readjusted after undershooting and overshooting. This is a situation that did not arise in Irwin's experiments; in his experiments, there was nothing more to read after the saccade, so there was no need to readjust the position of the eyes even if they had either undershot or overshoot.

In order to see whether this line of explanation is on the right track, the mean TR of each experimental condition was recalculated, this time excluding those trials in which the subject's eyes either overshoot or undershot the spill-over word. The result of this recalculation is shown below.



**Fig.4** Mean post-saccadic processing time (TR) in each condition, excluding trials involving under- or overshooting

As Fig. 4 shows, when trials involving under- or overshooting were excluded, the mean TR was shorter in the long-saccade condition than in the short-saccade condition, both in the high-frequency target word condition ( $289 < 294$ ) and in the low-frequency target word condition ( $331 < 357$ ). This lends support to the interpretation that I have been arguing for. Furthermore, the difference between the high-frequency condition and the low-frequency condition was smaller in the long-saccade condition ( $331 - 289 = 42\text{ms}$ ) than it was in the short-saccade condition ( $357 - 294 = 63\text{ms}$ ), even after trials involving under- or overshooting were excluded. This is consistent with the view that the interaction effect that was detected using the more inclusive data set was not an epiphenomenon arising from a particular distribution of errors in each experimental condition.

Should we have excluded from analysis those trials involving undershooting or overshooting? There were two reasons why we did not do so. First, it is not necessarily clear whether undershooting and overshooting are merely mistakes that are followed by a corrective saccade or word processing is initiated nearly normally even when the eyes have undershot or overshot. Second, there were unfortunately not a sufficient number of trials in which the eyes landed precisely on the spill-over word; it was not possible to exclude all trials involving undershooting or overshooting and still obtain the stable mean of each dependent variable.

It might appear that the reason why undershooting and overshooting took place so often must be that the saccade target box was too inconspicuous in the current experiment. However, a pilot study had shown that, even if the target box were more conspicuous, the error rate would not have been improved.

In the pilot study, two female subjects and one male subject, none of whom participated in the main experiment, read 150 sentences using saccade target boxes of three different sizes. The subjects read 50 sentences using the same target box as was used in the main experiment; a box which was 0.40 deg wide and 0.51 deg high was used as the target box in this condition (to be called the small box condition). The subjects read another 50 sentences using a somewhat more conspicuous target box; in this condition (to be called the medium box condition), the target box consisted of a box identical to that used in the small box condition and another, larger box whose inside was 1.19 deg wide and 1.51 deg high; the width of the lines that made up the larger box was 0.2 deg, and the smaller box was at the center of the larger box. And the subjects read another 50 sentences using an even more conspicuous target box; in this condition (to be called the large box condition) the target box consisted of a box identical to that used in the small box condition and another, larger box whose inside was 3.6 deg wide and 1.5 deg high; the width of the lines that made up the larger box was 0.1 deg, and the left-hand side of the smaller box was separated from the left-hand side of the larger box by 0.4 degrees, and the top of the smaller box was separated from the top of the larger box by 0.4 degrees. The order of the presentation of 150 sentences was randomized. The other procedures were the same with the current experiment. A trial was classified as 'acceptable' if the subject's eyes landed between the point that was one character space to the left of left-hand side of the box (the smaller box in the medium and the large box condition) and the point that was four character spaces to right of the left-hand side of the box. The percentage of 'acceptable' trials was larger in the small box condition than in the medium box condition or the large box condition.

Thus, I conclude that there was not an obvious flaw in the design of the experiment and that the fact that TR was longer in the long-saccade condition is merely a result of there being more undershooting and overshooting in the long-saccade condition and thus does not affect the conclusion that linguistic processing was not completely suppressed during saccades.

The third, and final complication to be discussed in this context is the fact that the subjects fixated the target words longer in the short-saccade condition than in the long-saccade condition, as shown in Fig. 3. This fact does not invalidate the main conclusion of this study; in fact, it strengthens it. If the subjects had fixated the target words longer in the long saccade condition than in the short saccade condition, then that would have weakened our case significantly; the reduction of the spill-over effect in the long-saccade condition could have been merely a consequence of the subjects' having spent more time processing the target words while fixating them. However, since the subjects in fact fixated the target words longer in the short saccade condition, the reduction of the spill-over effect in the long saccade condition cannot be ascribed to such an effect; rather, the reduction took place *despite* such an



effect, and therefore must be interpreted as evidence that cognitive processing was not suppressed completely during saccades.

To summarize, none of the three complications that have been examined invalidate the conclusion that linguistic processing was not completely suppressed during saccades while subjects read sentences.

What kind of linguistic processing is it that was not completely suppressed and continued even during saccades? This question is easy to answer in the case of Irwin's (1998) experiments; in his experiments, what continued during saccades is lexical access to each isolated word. However, in the current study, it is difficult to answer this question with much confidence.

Some serial models of eye movement control propose that an earlier stage of processing is a stage of orthographic identification that is influenced by word frequency and predictability (e.g. Rayner, Reichle, & Pollatsek, 2000; Reichle, Pollatsek, Fisher, & Rayner, 1998). However, other researchers suggest that processing beyond orthographic identification also influences initial fixation durations; from the early stage, the higher level processing such as full lexical access, syntactic processing, and thematic and semantic processing may begin and influence the course of processing (Liversedge & White, 2003).

The main problem here is that the source of the spill-over effects is not yet known. They might reflect the difficulty of integrating a low-frequency word into the evolving sentential representation, they might reflect the eye-mind lag, and there are other conceivable possibilities as well.

## 5.2. *Future directions.*

In this final subsection, I will briefly consider what future directions one might pursue as continuation of the present line of inquiry.

### 5.2.1. *Leftward saccades.*

One variant of the present study that might be worth considering is an experiment involving saccades to the left. For one thing, eye movements are not always from left to right even when the reader is reading sentences in English in a natural setting. Sometimes readers make right-to-left eye movements called regressions to fixate an earlier word that they would like to look at again, and at the end of every line, readers make a long leftward saccade called a return sweep, possibly followed by small corrective saccades further to the left (Carreiras & Clifton, Jr., 2004; Rayner, 1998). Thus if we are to study what happens during saccades when people are reading sentences, it seems reasonable to include the direction of saccade as

an independent variable. Moreover, experiments that use right-to-left saccades may generate results that are qualitatively different from those generated by experiments using left-to-right saccades. The saccade generation exhibits hemispheric asymmetries that depend on the direction of saccades involved; rightward saccades and leftward saccades are controlled by different cortical hemispheres (Brockmole, Carlson, & Irwin, 2002). This leads us to expect that saccades to the left might have different properties than saccades to the right; one might expect different types of saccadic suppression, depending on the direction of the saccade.

### 5.2.2. *Other types of linguistic material.*

In the present experiment, all the target words were common nouns. Thus, what was studied in the present study only concerns whether cognitive processing of common nouns is suppressed during saccades. One might envisage experiments that would test whether cognitive processing of other types of linguistic material is suppressed during saccades. For example, an experiment using pronouns as target words might be worth considering. When a reader encounters a pronoun, he or she recalls not just the content of its antecedent but also the location in the text (or whatever is being looked at) at which the antecedent had been seen. This has led Fisher to argue that reading requires spatial processing (Fisher, 1999). Thus, it might be possible to study whether spatial processing is suppressed during saccades by using pronouns as target words (Richard Shillcock, personal communication). More generally, it may be possible to gain more detailed information as to what types of cognitive processing is suppressed during saccades if we use various types of target words and various types of syntactic contexts for those words.

## 6. *Summary*

In this study, we set out to study whether cognitive processing is suppressed during saccades while people are reading sentences, as opposed to isolated words. We conducted an eye-tracking experiment in which subjects made either a long saccade or a short saccade after reading either a high-frequency word or a low-frequency word, and we measured, among other things, the time that elapsed between the onset of the saccade and the end of the fixation on the word immediately following the saccade location. It was expected that, if and only if cognitive processing was not completely suppressed during saccades, there should be an interaction effect between the corpus frequency of the pre-saccade word and the saccade length on the measured duration of time. And there was in fact such an interaction effect, which was significant in the ANOVA with the subjects as the random factor. Several

complicating factors were discussed and found not to invalidate the original hypothesis. Thus it was concluded that linguistic processing was not completely suppressed during saccades while people read sentences.

## 7. Acknowledgments

I am grateful to Martin Pickering, Scott McDonald, Martin Corley and Richard Shillcock for their invaluable suggestions based on their profound knowledge and experience throughout this project. A preliminary version of this work was presented at the 2<sup>nd</sup> Edinburgh Psycholinguistic Postgraduate conference, held in Edinburgh, and the 21<sup>st</sup> Psypag annual conference, held in Reading. I would like to thank the audience at those occasions for helpful comments.

## 8. References

- Balota, D. A., & Chumbley, J. J. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 340-357.
- Balota, D. A., Paul, S. T., & Spieler, D. H. (1999). Attentional control of lexical processing pathways during word recognition and reading. In S. Garrod & M.J. Pickering (Eds.), *Language processing* (pp. 15-57). East Sussex: Psychology Press.
- Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, 17, 364-390.
- Brockmole, J.R., Carlson, L.A., & Irwin, D.E. (2002). Inhibition of attended processing during saccadic eye movements. *Perception & Psychophysics*, 64, 867-881.
- Brown, G.D.A. (1984). A frequency count of 190,000 words in the London-Lund Corpus of English Conversation. *Behavior research methods, instruments, & computers*, 16, 502-532.
- Carreiras, M., & Clifton, C., Jr. (2004). On the on-line study of language comprehension. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension* (pp. 1-14). East Sussex: Psychology Press.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.
- Ehrlich, K., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20, 641-655.

- Fischer, M.H. (1999). Memory for word locations in reading. *Memory*, 7, 79-116.
- Gilhooly, K.J., & Logie, R.H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behaviour Research Methods and Instrumentation*, 12, 395-427.
- Glushko, R.J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674-691.
- Grishman, R., Macleod, C., & Meyers, A. (1994). Complex Syntax: Building a Computational Lexicon. *Proceedings of Computational Linguistics (COLING) 1994*, 268-272.
- Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P., & Grady, C.L. (1994). The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. *Journal of Neuroscience*, 14, 6336-6353.
- Hill, R.L., & Murray, W.S. (2000). Commas and Space: Effects of punctuation on eye movements and sentence parsing. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp. 565-590). Oxford: Elsevier.
- Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29-54). Oxford: Elsevier.
- Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40, 431-439.
- Irwin, D.E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology*, 36, 1-27.
- Irwin, D.E. (2003). Eye movements and visual cognitive suppression. *Psychology of Learning and Motivation*, 42, 265-293.
- Irwin, D.E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The Interface of Language, Vision, and Action*. East Sussex: Psychology Press.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Kapoula, Z., & Robinson, D. A. (1986). Saccadic undershoot is not inevitable: saccades can be accurate. *Vision Research*, 26, 735-743.
- Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, 45, 153-168.
- Kucera, H., & Francis, W.N. (1967). *Computational Analysis of Present-Day American English*. Providence: Brown University Press.

- Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: the influence of past, present, and future words on fixation duration. *Journal of Experimental Psychology: General*, 135, 12-35.
- Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (pp. 55-75). Oxford: Elsevier.
- Liversedge, S. P., & White, S. J. (2003). Psycholinguistic processes affect fixation durations and orthographic information affects fixation locations: Can E-Z Reader cope? *Behavioral and Brain Sciences*, 26, 492-493.
- Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? *Vision Research*, 44, 1013-1024.
- Mitchell, D.C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension* (pp. 15-32). East Sussex: Psychology Press.
- Mellet, E., Tzourio, N., Crivello, F., Jolio, M., Denis, M., & Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *Journal of Neuroscience*, 16, 6504-6512.
- Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neuroscience*, 6, 414-417.
- Murray, W.S. (2000). Sentence Processing: Issues and Measures. In A. Kennedy, R. Racach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 649-666). Oxford: Elsevier.
- Pickering, M.J., Frisson, S., McElree, B., & Traxler, M.J. (2004). Eye movements and semantic composition. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension* (pp. 33-50). East Sussex: Psychology Press.
- Paivio, A., Yuille, J. C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, 76, 1-25.
- Potter, M. (1984). Rapid serial visual presentation (RSVP): A method for studying language processing. In D. Kieras & M. Just (Eds.), *New methods in reading comprehension research* (pp. 91-118). Hillsdale, NJ: Erlbaum.
- Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology*, 16, 3-26.

- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14, 191-201.
- Rayner, K., Kambe, G., & Duffy, S.A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology*, 53A, 1061-1080.
- Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research*, 38, 1129-1144.
- Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: cognitive control of eye movements. *Psychological Science*, 14, 385-388.
- Rayner, K. & McConkie, G.W. (1976). What guides a reader's eye movements? *Vision Research*, 16, 829-837.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Hillsdale, NJ: Erlbaum.
- Rayner, K., Reichle, E.D., & Pollatsek, A. (2000). Eye movement control in reading: Updating the E-Z Reader model to account for initial fixation locations and refixations. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 701-720). Oxford: Elsevier.
- Rayner, K., & Sereno, S. C. (1994). Eye movements in reading: Psycholinguistic studies. In Gernsbacher, M. A. (Ed.), *The handbook of psycholinguistics* (pp. 57-81). London: Academic Press.
- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4, 21-49.
- Rayner, K., Sereno, S. C., & Raney, G.E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1188-1200.
- Rayner, K., & Well, A., (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3, 504-509.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105, 125-157.
- Sanders, A. F. (1963). *The selective process in the functional visual field*. Assen, The Netherlands: Van Gorcum.

- Sanders, A. F. (1970). Some aspects of the selective process in the functional visual field. *Ergonomics*, 13, 101–117.
- Schall, J.D., Hanes, D.P., Thompson, K.G., & King, D.J. (1995). Saccade target selection in frontal eye field of macaque. I. Visual and premovement activation. *Journal of Neuroscience*, 15, 6905–6918.
- Schall, J.D., Morel, A., King, D.J., & Bullier, J. (1995). Topography of visual cortical afferents to frontal eye field in macaque: Functional convergence and segregation of processing streams. *Journal of Neuroscience*, 15, 4464–4487.
- Shilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26, 1270–1281.
- Stanton, G.B., Bruce, C.J., & Goldberg, M.E. (1993). Topography of projections to the frontal lobe from the macaque frontal eye fields. *Journal of Comparative Neurology*, 330, 286–301.
- Takahashi, K., Nakayama, M., & Shimizu Y. (2001). The eye-movement and pupil size respond to calculation task in viewing a target. *The Transaction of the Institute of Electronics, Information and Communication Engineers*, J84-D-I, 1573–1576.
- Tanenhaus, M.K. (2004). On-line sentence processing: past, present and, future. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension* (pp. 371–392). East Sussex: Psychology Press.
- Toglia, M.P., & Battig, W.R. (1978). *Handbook of Semantic Word Norms*. New York: Erlbaum.
- Ungerleider, L.G., & Haxby, J.V. (1994). ‘What’ and ‘where’ in the human brain. *Current Opinion in Neurobiology*, 4, 157–165.
- Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (Eds.), *Analysis of Visual Behavior*. Cambridge, Mass.: MIT Press.
- Van Duren, L., & Sanders, A. F. (1995). Signal processing during and across saccades. *Acta Psychologica*, 89, 121–147.
- Vitu, F. (1991). The influence of parafoveal processing and linguistic context on the optimal landing position effect. *Perception & Psychophysics*, 50, 58–75.
- Vonk, W., & Cozijn, R. (2003). On the treatment of saccades and regressions in eye movement measures of reading time. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: cognitive and applied aspects of eye movement research* (pp. 291–312). Oxford: Elsevier.

Wilson, M.D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2, *Behavioural Research Methods, Instruments and Computers*, 20, 6-11.



## Appendix.

In the following sentences, the symbol “>>” indicates the saccade location. The odd-numbered sentences contain high-frequency target words and the even-numbered sentences contain low-frequency target words.

1	The young officer	>>	entered the hut.
2	The young samurai	>>	entered the hut.
3	The angry mother	>>	ignored the food.
4	The angry barman	>>	ignored the food.
5	The tired student	>>	left the plane.
6	The tired mourner	>>	left the plane.
7	The aged lawyer	>>	calmed the child.
8	The aged oilman	>>	calmed the child.
9	A very old window	>>	smashed to bits.
10	A very old beaker	>>	smashed to bits.
11	The broken device	>>	upset them all.
12	The broken peeler	>>	upset them all.
13	The sick minister	>>	kept on talking.
14	The sick polyglot	>>	kept on talking.
15	The dirty worker	>>	crossed the road.
16	The dirty jackal	>>	crossed the road.
17	The shiny express	>>	moved very fast.
18	The shiny frigate	>>	moved very fast.
19	The shaky vehicle	>>	creaked loudly.
20	The shaky omnibus	>>	creaked loudly.
21	The large prison	>>	remained empty.
22	The large hangar	>>	remained empty.
23	The older system	>>	functions better.
24	The older glider	>>	functions better.
25	The noisy brother	>>	chased the swan.
26	The noisy moorhen	>>	chased the swan.
27	The plump husband	>>	found the book.
28	The plump harpist	>>	found the book.
29	The sandy surface	>>	looked bumpier.
30	The sandy anthill	>>	looked bumpier.
31	The dirty island	>>	became an issue.
32	The dirty cutlet	>>	became an issue.
33	The young cousin	>>	annoyed our son.
34	The young weasel	>>	annoyed our son.
35	The proud author	>>	gave up the job.

36	The proud lodger	>>	gave up the job.
37	The bluish object	>>	caught my eye.
38	The bluish maggot	>>	caught my eye.
39	The horrid action	>>	angered my son.
40	The horrid geezer	>>	angered my son.
41	Too much freedom	>>	causes trouble.
42	Too much anthrax	>>	causes trouble.
43	Its sheer luxury	>>	amazed the man.
44	Its sheer idiocy	>>	amazed the man.
45	The bland coffee	>>	would upset them.
46	The bland fillet	>>	would upset them.
47	My careful nature	>>	reassured them.
48	My careful rigour	>>	reassured them.
49	Too much success	>>	repels some boys.
50	Too much conceit	>>	repels some boys.
51	Hiding the number	>>	would be stupid.
52	Hiding the sachet	>>	would be stupid.
53	The honest driver	>>	paid the price.
54	The honest jester	>>	paid the price.
55	The rich company	>>	should help her.
56	The rich insurer	>>	should help her.
57	His magic method	>>	might save them.
58	His magic locket	>>	might save them.
59	The daily report	>>	upset the girl.
60	The daily tumult	>>	upset the girl.
61	I crammed pencils	>>	into our sack.
62	I crammed syphons	>>	into our sack.
63	I piled magazines	>>	onto the table.
64	I piled stilettos	>>	onto the table.
65	I put the picture	>>	onto the wall.
66	I put the wrapper	>>	onto the wall.
67	Finding a market	>>	might be harder.
68	Finding a geyser	>>	might be harder.
69	I pinned a notice	>>	above the desk.
70	I pinned a toupee	>>	above the desk.
71	I flung a journal	>>	onto the ground.
72	I flung a bassoon	>>	onto the ground.
73	He tucked cotton	>>	around the cat.
74	He tucked muslin	>>	around the cat.

75	He ferried cattle	>>	across the river.
76	He ferried seamen	>>	across the river.
77	I put my patient	>>	under the roof.
78	I put my spaniel	>>	under the roof.
79	I leaned weapons	>>	against the wall.
80	I leaned valises	>>	against the wall.
81	Both the members	>>	backed the plan.
82	Both the vandals	>>	backed the plan.
83	All the teachers	>>	hate the mayor.
84	All the florists	>>	hate the mayor.
85	A certain problem	>>	nauseated her.
86	A certain trainee	>>	nauseated her.
87	All those answers	>>	appeared benign.
88	All those looters	>>	appeared benign.
89	The whole debate	>>	excited the kid.
90	The whole fracas	>>	excited the kid.
91	Each new mystery	>>	amused us a bit.
92	Each new poetess	>>	amused us a bit.
93	Maths and history	>>	interested them.
94	Maths and gunnery	>>	interested them.
95	Peas and chicken	>>	disgusted her.
96	Peas and sirloin	>>	disgusted her.
97	Some old drivers	>>	read the letter.
98	Some old welders	>>	read the letter.
99	Ed and the leader	>>	became friends.
100	Ed and the mugger	>>	became friends.
101	Then the children	>>	went to the city.
102	Then the herdsmen	>>	went to the city.
103	So the merchant	>>	fell fast asleep.
104	So the clansman	>>	fell fast asleep.
105	Oddly the police	>>	liked the man.
106	Oddly the papist	>>	liked the man.
107	Soon the audience	>>	will stand up.
108	Soon the helmsmen	>>	will stand up.
109	Here his language	>>	seemed too nice.
110	Here his lenience	>>	seemed too nice.
111	Often the machine	>>	bothered them.
112	Often the clutter	>>	bothered them.
113	Later the speaker	>>	went shopping.
114	Later the tipster	>>	went shopping.

115	Now his daughter	>>	seems very tall.
116	Now his godchild	>>	seems very tall.
117	Maybe the office	>>	depressed him.
118	Maybe the fizzle	>>	depressed him.
119	There the people	>>	looked so awful.
120	There the gibbon	>>	looked so awful.
121	I placed blankets	>>	under the table.
122	I placed jodhpurs	>>	under the table.
123	Soon the animals	>>	were dying out.
124	Soon the cicadas	>>	were dying out.
125	Surely her salary	>>	must bother her.
126	Surely her eczema	>>	must bother her.
127	Now this entrance	>>	looks dangerous.
128	Now this flywheel	>>	looks dangerous.
129	The giant father	>>	scared her cat.
130	The giant baboon	>>	scared her cat.
131	A tall engineer	>>	started shouting.
132	A tall exorcist	>>	started shouting.